DOCUMENT RESUME

ED 468 488 TM 034 398

AUTHOR	Henderson, Dianne L.; Julian, Marc W.; Yen, Wendy M.
TITLE	Multiple Measures: Examination of Alternative Design and Analysis Models.
PUB DATE	2002-04-00
NOTE	20p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (New Orleans, LA, April 2-4, 2002).
PUB TYPE	Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	Academic Achievement; *Educational Assessment; Educational Change; Elementary Secondary Education; Evaluation Methods; *Measurement Techniques; *Research Design; Statistical Analysis
IDENTIFIERS	*Multiple Measures Approach; National Assessment of Educational Progress; No Child Left Behind Act 2001

ABSTRACT

This paper discusses different perspectives associated with the use of multiple measures in educational assessment and explores some of the technical considerations for selecting and combining multiple measures. The paper concludes with an example of the use of multiple measures with regard to the recent No Child Left Behind (NCLB) Act of 2002. "Multiple measures" can refer to a variety of possible scenarios that combine multiple sources of information within educational assessment, but, in general, multiple measures means the use of more than one assessment measure to evaluate student performance. A multiple measures approach attempts to increase reliability by decreasing the amount of error associated with evaluation at all levels. How to combine the information is one of the most difficult aspects of using multiple measures. The National Assessment of Educational Progress (NAEP) can be used as a multiple measure, especially as an independent measure to confirm the validity of conclusions reached about student achievement in the context of the NCLB Act. The National Assessment Governing Board has acknowledged the complexity of using NAEP results in a confirming role for the NCLB, but has been exploring ways to use NAEP results as confirmatory evidence for NCLB evaluation. (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 468 488

Running head: MULTIPLE MEASURES

Multiple Measures:

Examination of Alternative Design and Analysis Models

Dianne L. Henderson

Marc W. Julian

Wendy M. Yen

ETS K-12 Works, Inc.[1]

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

D. Henderson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

TM034398

[1] The opinions expressed in this paper are those of the authors and not necessarily those of ETS K-12 Works, Inc.

Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, LA, April, 2002.

Multiple Measures:

Examination of Alternative Design and Analysis Models

The call for the use of multiple measures has become increasingly pervasive within today's educational climate. The use of multiple measures is stressed by legislation regulating the distribution of Title 1 funding to states and schools, as well as professional and industry standards regarding the use of test scores for high stakes decisions. However, despite widespread agreement about the need to use multiple measures for making decisions about individuals and institutions, there is no clear consensus on what constitutes an acceptable measure, the minimal technical criteria required for each measure, the methods of synthesizing and using information from multiple sources, and effective procedures for communicating this information to the public. This paper discusses different perspectives associated with the use of multiple measures, as well as some of the technical considerations for selecting and combining multiple measures. The paper will conclude with an example of the use of multiple measures with regard to the recent No Child Left Behind (NCLB) legislation (The No Child Left Behind Act of 2002, P. L., 107-110).

What are Multiple Measures?

The increasing call for multiple measures in assessment is correlated with increasing pressure for individual and organizational accountability. Assessments that carry high stakes are more prevalent than ever before for individual students, teachers, schools, and states. Multiple measures are intended to improve the quality of high stakes decision, but the definition of multiple measures and how they should be used in high stakes decisions are not clear.

Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, LA, April, 2002.

Legislative requirements, such as NCLB, are a driving force behind the focus on the use of multiple measures to evaluate performance at the school, local educational agency (LEA), and state levels. In order to receive Title 1 funding, each state must demonstrate the development or adoption of a set of high-quality yearly student assessments to be used to evaluate the yearly performance of each LEA and school. In addition, the NCLB Act requires states to include multiple measures of student academic performance to determine yearly state performance.

Professional and industry standards also speak directly to the need for multiple measures when high stakes decisions are being made at the individual student level. Multiple measures are explicitly recommended within the "Standards for Educational and Psychological Testing (AERA/APA/NCME, 1999). Standard 13.7 states:

> In educational settings, a decision or characterization that will have major impact on a student should not be made on a simple test score. Other relevant information should be taken into account if it will enhance the overall validity of the decision (pp. 147-148).

In alignment with the Standards, the American Educational Research Association (AERA, July, 2000) provides additional support for the use of multiple measures through the following position statement:

> ... decisions that affect individual students' life chances or educational opportunities should not be made on the basis of test scores alone. Other relevant information should be taken into account to enhance the overall validity of such decisions (p. 1).

The National Research Council report *High Stakes* (1995) also stresses that multiple assessments are critical to making appropriate decisions and inferences about students, teachers, and schools. In addition, the testing and assessment industry also encourage the use of multiple measures for high stakes decisions about individuals. For example, the *ETS Standards for Quality and Fairness* (Educational Testing Service, 2000) state that "...assessment users ... need to allow students who must pass a test to be promoted or granted a diploma, or individuals seeking certification, a reasonable number of opportunities to succeed" (p. 58).

As described above, the interpretation of the actual definition of multiple measures can vary. Based on these examples it is evident that multiple measures can refer to an enormous variety of possible scenarios that combine multiple sources of information within educational assessment. As the professional and industry standards directly state, students should have multiple opportunities to take tests; especially when high stakes decisions are being made about students. However, the opportunity for a student to retake a test multiple times is one of the most limited definitions of multiple measures. More general definitions of multiple measures can also refer to the premise that decisions made about the student should be based on multiple sources of information.

In more general terms, multiple measures encompass the use of more than one assessment measure within the same or different content areas (e.g., Reading, Mathematics); the use of different types of assessment information (e.g., norm-referenced, criterion-referenced, performance assessments, teacher-made test scores, and classroom portfolio assessment); or the use of different item types (e.g., selected-response (SR), constructed-response (CR), performance assessments (PA), and portfolios).

Moreover, multiple measures can also include non-cognitive measures such as attendance or school and community participation.

In addition, the definition of multiple measures can include the administration of an assessment over multiple occasions or across multiple cohorts to obtain longitudinal information encompassing multiple levels of analysis. That is, individual scores can be aggregated together to evaluate the performance at the classroom, school, LEA, and state levels. This summary information can also be disaggregated to evaluate performance of various subgroups at the classroom, school, LEA, and State level. Finally, a multiple measures approach can also encompass multiple sources of information for various levels of analyses. That is, a multiple measures approach may also refer to the assessment of individual and school performance using performance on one or more achievement assessments in conjunction with other non-cognitive indicators such as student attendance, for evaluation of the current performance and for comparisons to previous performance.

Using the most inclusive definition of multiple measures, information from multiple sources (instruments, methods, occasions, cohorts and/or levels) is gathered and synthesized in order to inform decisions about students, schools, and LEAs. By this definition, there is no single assessment instrument, whether purchased off the shelf or customized for a particular school, LEA, or state, that meets these requirements. Attending to the multiple goals of a testing program using multiple instruments undoubtedly increases the complexity of the program, however, the guiding principle for implementing a multiple measures approach for an assessment system is to improve the reliability and validity of educational decisions.

Validity and Reliability

A multiple measures approach attempts to increase reliability by decreasing the amount of error associated with the evaluation at all levels, including students, teachers, schools, and LEAs. Assessment systems that incorporate multiple measures are more likely to be able to address complex accountability questions about students, teachers, and schools because they have multiple sources of information to evaluate. The inclusion of different types of items (SR, CR, PA) can provide a more complete picture of a given construct as well as individual achievement by providing multiple ways for students to demonstrate knowledge and mastery of the content area. Multiple measures can also provide a more comprehensive-alignment between assessment systems and the state content and performance standards.

While in the ideal case, multiple measures can be used to gather comprehensive, corroborating evidence of performance, it is also possible that additional measures of academic and school performance can provide divergent evidence of achievement. Therefore, it is important to understand the goals and purposes of not only the assessment system, but also each source of information available and how the multiple sources of information can be used to make more accurate high stakes decisions.

Incorporating Information From Multiple Measures for Multiple Levels of Analysis

One of the more complex challenges associated with the use of multiple measures is the need to synthesize all the data in an efficient, useful, understandable, and legally defensible manner. Information obtained from multiple sources has to be relevant to the final evaluations and subsequent high-stakes decisions that will be made about individuals, teachers, and schools. Given that information obtained from different sources

are likely to be used at more than one aggregate level, it is critical that the different measures and the associated outcomes provide sufficient information to ensure the information needs of all reporting levels are addressed

The selection of measures for use in a state assessment program needs to be evaluated with respect to the multilevel goals of the assessment system. In determining which sources of information to include, the state should ensure that sufficient information is obtained from the various sources of information to evaluate student, classroom and school performance. Questions that can help focus this evaluation include,

- What is the primary purpose of each assessment?
- How will information be used at each level of analysis?
- Does each assessment have sufficient reliability for reporting individual scores?
- How will assessment performance this information be aggregated to evaluate classrooms/teachers and/or schools?
- What information do the measures provide for each level of analyses (student, classroom, school) and what information is missing?
- Could additional measures be added to provide better coverage?

Combining Information from Multiple Measures

One of the most difficult aspects of using a multiple measures approach is determining how to combine the information. There are four primary ways in which information from multiple measures can be used: conjunctively, compensatory, mixed conjunctive-compensatory, and confirmatory.

A conjunctive approach to integrating information across multiple measures requires the demonstration of a minimum level of performance across all measures. In contrast, in the compensatory approach, poor performance on one or more measures can be counterbalanced by higher performance on another measure. The mixed approach uses a combination of conjunctive and compensatory approaches. For example, a minimal performance level is required across measures, but beyond the minimal level of performance, poorer performance on one measure can be counterbalanced by better performance on other measures.

The fourth approach, confirmatory, is one in which information from one measure is used to confirm or compare information from another independent measure or measures. This can be a useful approach, however, it is important to consider whether or not it is psychometrically appropriate to use these independent measures as confirmation of performance and how to interpret the similarities or differences observed between the two measures. It is also important to recognize that a single measure of performance may be used to evaluate a more complex composite performance based on multiple indicators. For example, the NCLB Act creates the situation in which NAEP will be used in a confirming role for state assessments, which might be consist of more than one measure of achievement. This example is discussed in more detail in the next section.

No Child Left Behind and NAEP:

An Example of the Use of Multiple Measures in a Confirming Role

Recent No Child Left Behind (NCLB) federal legislation (The No Child Left Behind Act of 2002, P. L. 107-110) mentions the use of "multiple measures," as in the following example:

> Each State plan shall demonstrate that the State educational agency…has implemented a set of high-quality, yearly student academic assessments that include, at a minimum, academic assessments in mathematics, reading or language arts, and science that will be used as the primary means of determining the yearly performance of the State … in enabling all children to meet the State's challenging student academic achievement standards…. Such assessments shall …. involve multiple up-to-date measures of student academic achievement…. (pp. 25-26).

Further definitions or requirements for the "multiple measures" are not provided in the legislation. However, the NCLB Act has created circumstances where the National Assessment of Educational Progress (NAEP) will be used in the role of "confirming" the results of state assessments in reading and mathematics. That is, NAEP will be used as a multiple measure. The following discussion examines this circumstance in more detail to exemplify the considerations that can be entailed in using multiple measures in a confirming role.

Legislation and Policy

While the NCLB does not prescribe the use of NAEP in a confirming role for state assessments, that apparently was the legislative and executive intent. A recent report published by the National Assessment Governing Board (NAGB; Ad Hoc Committee on Confirming Test Results, 2002) describes the situation:

> Based on the legislation and official statements contained in the respective fact sheets, the following conclusions are drawn about the role of the National Assessment under the No Child Left Behind Act:

- the National Assessment must conduct biennial assessment in reading and mathematics in grades 4 and 8 as its first priority
- participation of states is required in biennial assessments in reading and mathematics in grades 4 and 8 conducted by the National Assessment, beginning in school year 2002-2003
- participation in such assessments is required of schools that receive Title I funding and are selected for the NAEP sample
- NAEP state results in grades 4 and 8 reading and mathematics may be used to "confirm" state test results under the purview of the U. S. Department of Education
- NAEP state results in grades 4 and 8 reading and mathematics will not be used to make awards to or exact penalties against states. (p. 3)

Regardless of whether the U.S. Department of Education conducts confirming analyses, the biennial NAEP data will be available for all states wishing to receive federal Title I funding. These NAEP results will be analyzed by the states themselves, by the media, and/or by other interested parties. In other words, the simple availability of the NAEP data will assure that multiple measures of state performance will exist and someone will be comparing state and NAEP results.

Motivation for Using NAEP Results as a Multiple Measure in a Confirming Role

The motivation for using NAEP as an "independent Benchmark" is described in a quote in a House of Representatives Fact Sheet (December 2001):

> "While states need the flexibility to develop their own assessments, there must also be an external benchmark against which to compare the rigor of their standards, tests and accountability systems. The National Assessment of Educational Progress (NAEP) provides such a marker. . .We also [favor] strengthening NAEP's independence and taking additional steps to ensure that its tests are of high quality and its results speedily and accurately reported." (p. 5)

This Fact Sheet emphasizes "rigor of standards, tests and accountability systems." Such a statement could be interpreted to mean the challenge or difficulty of meeting the state standards. For example, a recent *Education Week* article ("What's Proficient?" February 20, 2002) compared the percents of students reaching each state's 4th and 8th grade mathematics Proficient standard to the percents in that state reaching the Proficient standard on the NAEP assessment. In the vast majority of cases, a lower percent of students were Proficient on NAEP.

Thus, one use of NAEP as a multiple measure would be expected to be an evaluation of the rigor of state standards at one point in time. In addition NAEP can be expected to be used to evaluate changes over time in student performance. In particular, the NCLB legislation requires that

> Each state shall establish a timeline for adequate yearly progress. The timeline shall ensure that not later than 12 years after the end of the 2001-2002 school year, all students in each group described in subparagraph (C)(v) will meet or exceed the State's proficient level of academic achievement on the State assessments.... (NCLB Act, pp. 23-24)

The referenced subgroups are "economically disadvantaged students; …students from major racial and ethnic groups; … students with disabilities; and … students with limited English proficiency" pp. 22-23). The NCLB Act is particularly concerned with "…closing the achievement gap between high- and low-performing children, especially the achievement gaps between minority and nonminority students, and between disadvantaged children and their more advantaged peers" (p. 16). In fact the complete title of this legislation is "An act to close the achievement gap with accountability, flexibility, and choice, so that no child is left behind" (p. 1). Thus, in addition to using NAEP as a multiple measure to monitor the rigor of state standards and assessments, it is to be expected that NAEP will be used to confirm growth in student achievement and the reduction of gaps in performance among subgroups of students.

Similar uses of NAEP have occurred in the past for individual states when questions have been raised about a testing program. For example, in 1995 at the request of the state legislature, an independent panel reviewed the Kentucky Instructional Results Information System (Hambleton, Jaeger, Koretz, Linn, Millman, & Phillips, 1995). One of the issues addressed was "Is there evidence that education has improved in Kentucky?" One set of evidence that was examined was changes over years in state NAEP results as compared with changes in state test results.

Thus, the primary policy goals in using NAEP as a multiple measure in NCLB is for an independent measure to confirm the validity of conclusions reached about student achievement:

- That the state standards and tests are sufficiently rigorous

- That progress over years in student achievement that is reflected in state test scores is also reflected in NAEP results
- That progress in closing achievement gaps seen with state test data is also reflected in NAEP scores

Differences Between NAEP and State Assessments that May Impact Their Use as Multiple Measures

This NCLB example makes it clear that in developing and evaluating models for using multiple measures, the specific application must be considered. For example, given that NAEP is to be used as an independent measure, issues of mathematically combining results from multiple measures and using compensatory or non-compensatory models are not relevant. Of greatest relevance is whether analysis of NAEP results and state test results lead to similar conclusions about student achievement. In creating a rational procedure for jointly considering the two sets of results, it is essential to consider differences in the design and measurement properties of a state test and NAEP and the effects that these differences can have on the conclusions reached. The major differences in these assessment systems relate to content coverage, item format and difficulty, test administration, sub-group definition and inclusion, and scoring.

Content. Each state's assessment is developed to measure state content standards. NAEP is designed to reflect a broad national consensus of what students should know and be able to do at grades 4, 8, and 12. In fact, NAEP is administered in a matrix sampling design in order to maintain a breadth of content coverage while minimizing the amount of testing required of any one student. Figure 1 presents a simple diagram of how the content coverage of a state test and NAEP might overlap. The amount of overlap, and

its relationship to instruction, can systematically affect the similarity of results found with a state test and NAEP. In general one would expect NAEP to be less responsive to local instruction than is the state assessment. Differences that are found would need to be evaluated for their relevance to conclusions about whether a state is appropriately improving student achievement.

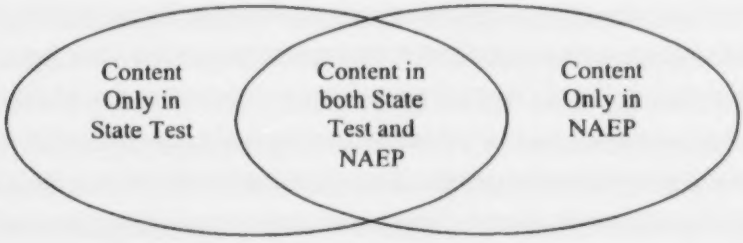

Figure 1. Hypothetical overlap of content coverage in a state test and NAEP

Item format and difficulty. A state test and NAEP can differ in terms of their proportional use of constructed-response versus multiple-choice items and in terms of the difficulty of the items included in the two assessments. Rubrics for evaluating constructed-response items can differ across the assessments.

Test administration. State assessments under NCLB will be administered annually on a census basis to students in grades 3 to 8, while NAEP will be administered biennially to a sample of students in grades 4 and 8. NAEP and state assessments can be administered at different times in the school year and therefore reflect somewhat different spans of instruction.

The stakes associated with an assessment are expected to affect teacher and student behavior with respect to test preparation and the level of motivation associated with test taking. State assessments are administered with high stakes, and these stakes are

bound to increase with the federal reward/sanction system and publicity added by NCLB. Under NCLB, all students will receive scores for their performance on the state test. Historically, NAEP has been administered under low stakes, although its NCLB confirming role will increase its stakes. With NAEP no student scores are calculated or returned to students.

Sub-group definition and inclusion. In evaluating the closure of gaps, it is necessary to identify students by subgroup. States can differ greatly in their definitions of disabilities and English language proficiency. While both state assessments and NAEP are intended to be inclusionary, they can differ in rules followed for student inclusion and test modifications or accommodations.

Scoring. NAEP reports estimated distributions of student scale scores, as well as estimated percents of students reaching NAEP proficiency levels (Advanced, Proficient, and Basic). States differ in terms of the proficiency levels that they have established and the percents of students achieving state proficiency levels versus the percents reaching the NAEP levels (n.b. the *Education Week* article, What's Proficient?"). In other words, the Proficient level for NAEP may be set at a different part of the achievement continuum than the Proficient level defined in a state. Depending on where instruction is focused in the state, the two assessment systems can reflect that instruction differently. For example, if instructional resources are particularly focused on low-achieving students and the state's Proficient level is set at a lower achievement level than is NAEP's Proficient level, improvements in student achievement may be reflected in increasing percents of Proficient students on the state test when that growth is not reflected on NAEP. (To minimize the misinterpretations that can occur from differing definitions of proficiency

levels, growth throughout the entire distribution of student scale scores be examined, not just percents of students at particular proficiency levels.)

Using NAEP Constructively as a Multiple Measure for NCLB

State tests and NAEP have many differences in their design and administration. These differences can contribute to differences in conclusions reached about student achievement under the NCLB. In some cases these differences are irrelevant to the goal of using NAEP as an independent measure of student achievement. For example, if differences in results are found because of different, but equally defensible, sub-group definitions, then NAEP's lack of confirmation deserves to be discounted. On the other hand, the differences can be quite relevant to the conclusions being reached. For example, if a state is showing impressive growth because of a narrow focus on simple content, then a lack of growth on NAEP's broader and more challenging content would be a notable finding. In some cases, it will be difficult to parse out the causes of differences between NAEP and state test results. For example, definitions of content domains can be complex and reasonable people can disagree about the importance of particular types of content. Differences in NAEP and state results that are attributable to differences in content coverage will deserve careful, non-superficial analysis.

The National Assessment Governing Board acknowledged the complexity of using NAEP results in a confirming role for NCLB. It established its Ad Hoc Committee on Confirming Test Results to consider this complexity and make recommendations about appropriate actions to be taken so that NAEP was prepared to do the best possible job in fulfilling its legislated role. In essence, this Committee conducted an in-depth examination of one particular situation involving multiple measures and attempted to

make constructive suggestions about how such multiple measures could be soundly implemented. In considering the NCLB circumstances, it became clear to the Ad Hoc Committee that states would be "making their cases" about state performance given all the evidence at their disposal. These cases would not be the result of the application of cut-and-dried formulas but informed and thoughtful analysis of individual state circumstances. "Thought experiments" were conducted with existing data for several states to hypothesize how the states might make their cases. Based on these examples and other discussions, the Ad Hoc Committee on Confirming Test Results made several recommendations to NAGB, for example:

> 2. "Informed judgment" and a "reasonable person" standard should be applied in using National Assessment data as confirmatory evidence for state results. Confirmation should not be conducted on a "point by point" basis or construed as a strict "validation" of the state's test results. (p. 8)
>
> 3. Limitations in using NAEP to confirm the general trend of state test results should be acknowledged explicitly. (p. 9)
>
> 7. In using NAEP achievement levels as confirmatory evidence, the percent at or above basic, proficient, and advanced, and the percent below basic, should always be presented and considered in light of the full range of state standards. Information about the achievement distribution should be used to augment the standards based interpretations of the results. (p. 12)

The full report of the Ad Hoc Committee on Confirming Test Results provides a valuable resource to those interested in exploring in-depth one example of the use of multiple measures.

Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, LA, April, 2002.

References

American Educational Research Association. (July, 2000). *AERA Position Statement Concerning High-Stakes Testing in PreK-12 Education* [electronic version]. Retrieved March 18, 2002, from http://www.aera.net/about/policy/stakes.htm

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing.* Washington, DC: American Psychological Association.

Ad Hoc Committee on Confirming Test Results. (March 1, 2002). *Using the National Assessment of Educational Progress to confirm state test results; A report of the Ad Hoc Committee on Confirming Test Results.* Washington, DC: National Assessment Governing Board.

Educational Testing Service. (2000). *ETS standards for quality and fairness.* Princeton, NJ: Author.

Fact Sheet on H. R. 1, Chairman John Boehner, December 10, 2001. Retrieved March 20, 2002 from House of Representatives website: http://edworkforce.house.gov/issues/107th/education/nclb/accountfact.htm

Hambleton, R. K., Jaeger, R. M., Koretz, D., Linn, R. L., Millman, J., & Phillips, S. E. (1995). *Review of the measurement quality of the Kentucky Instructional Results Information System, 1991-1994.*Office of Educational Accountability, Kentucky General Assembly.

What's Proficient? (2002, February 20), *Education Week.*

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

TM034398

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Multiple Measures: Examination of Alternative Design and Analysis Models

Author(s): Dianne L. Henderson, Marc W. Julian & Wendy M. Yen

Corporate Source: Educational Testing Services
presented at NCME, April, 2002

Publication Date: April 2002

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1

☑

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A

☐

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B

☐

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature: DHenderson	Printed Name/Position/Title: Dianne L. Henderson, Sr. Research Scientist
Organization/Address: Educational Testing Service, 60 Garden Court, Suite 202, Monterey, CA 93940	Telephone: 831 647 3745 / FAX: 831-647-3748
	E-Mail Address: dhenderson@ets.org / Date: Aug 30 02

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC CLEARINGHOUSE ON ASSESSMENT AND EVALUATION
UNIVERSITY OF MARYLAND
1129 SHRIVER LAB
COLLEGE PARK, MD 20742-5701
ATTN: ACQUISITIONS

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706

Telephone: 301-552-4200
Toll Free: 800-799-3742
FAX: 301-552-4700
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com

EFF-088 (Rev. 2/2000)